AGGREGATION TOOL TO CREATE *CURATED DATA* ALBUMS TO SUPPORT DISASTER RECOVERY AND RESPONSE

Rahul Ramachandran*, *GHRC NASA/MSFC*
rahul.ramachandran@nasa.gov

Ajinkya Kulkarni, Manil Maskey, Rohan Bakare,
Sabin Basyal, Xiang Li, Shannon Flynn
University of Alabama in Huntsville

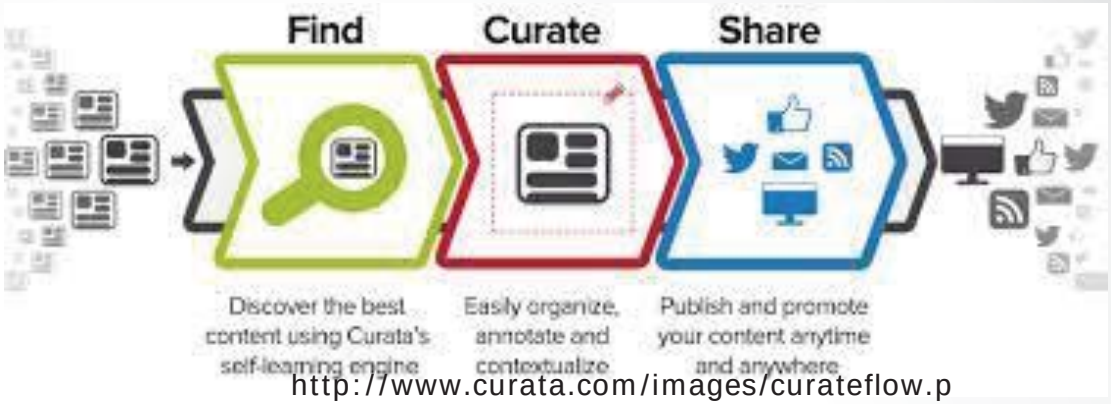
AGU Fall Meeting 2014 Session ID#: 2452

Helping Disasters Management through the Use of Remote Sensing Observational Data and their underlying Cyberinfrastructures

Outline

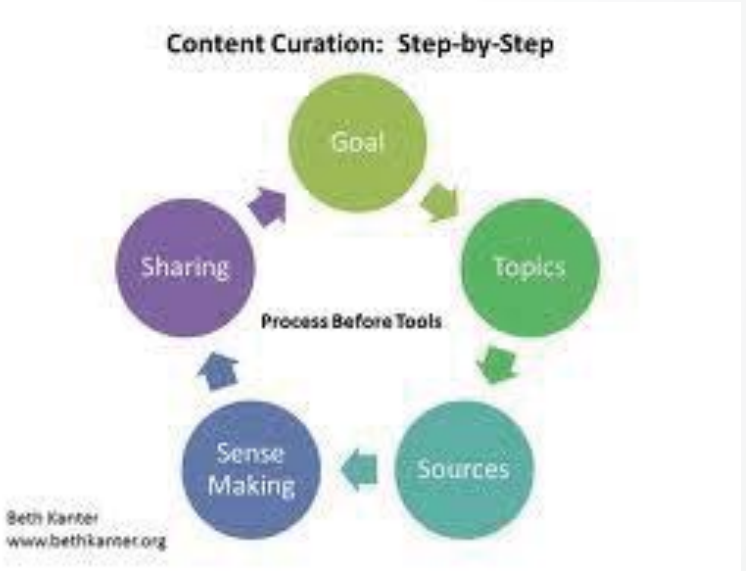
- Disaster Recovery and Response
  - Need for curated data and information
  - Emergency Management Process
- Data Albums
  - Information Seeking Models and Sensemaking
  - Architecture and technology overview
  - Data Album Instances Example
    - Severe Storms
    - Hurricane Case Studies
- Future work
  - Adapting Data Albums technology for Disaster Support

Curation

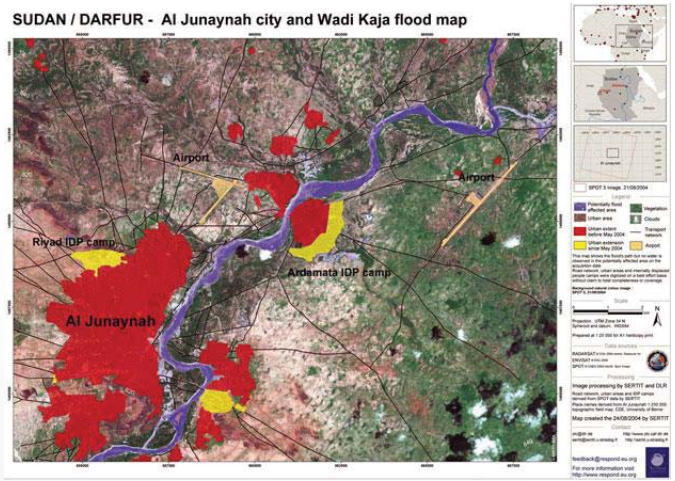


http://www.curata.com/images/curateflow.png

- The act of gathering, organizing and maintaining a collection of "resources" around a theme or a topic



http://farm7.static.flickr.com/6159/6212420184_b2bae4f5cb.jpg



http://www.eohandbook.com/eohb05/images/fig_05_(disaster).jpg

Resources: information + data

GHRC
Global Hydrology Resource Center

Role of Curated data: Situational Awareness Perspective

- Situational Awareness or Common Operating Picture – “describe getting real-time tactical information into the hands of the warfighter [or first responder] .. This information has to be part of *a seamless pipeline of data* which spans the breadth of permission planning, mission rehearsal, and mission execution.”

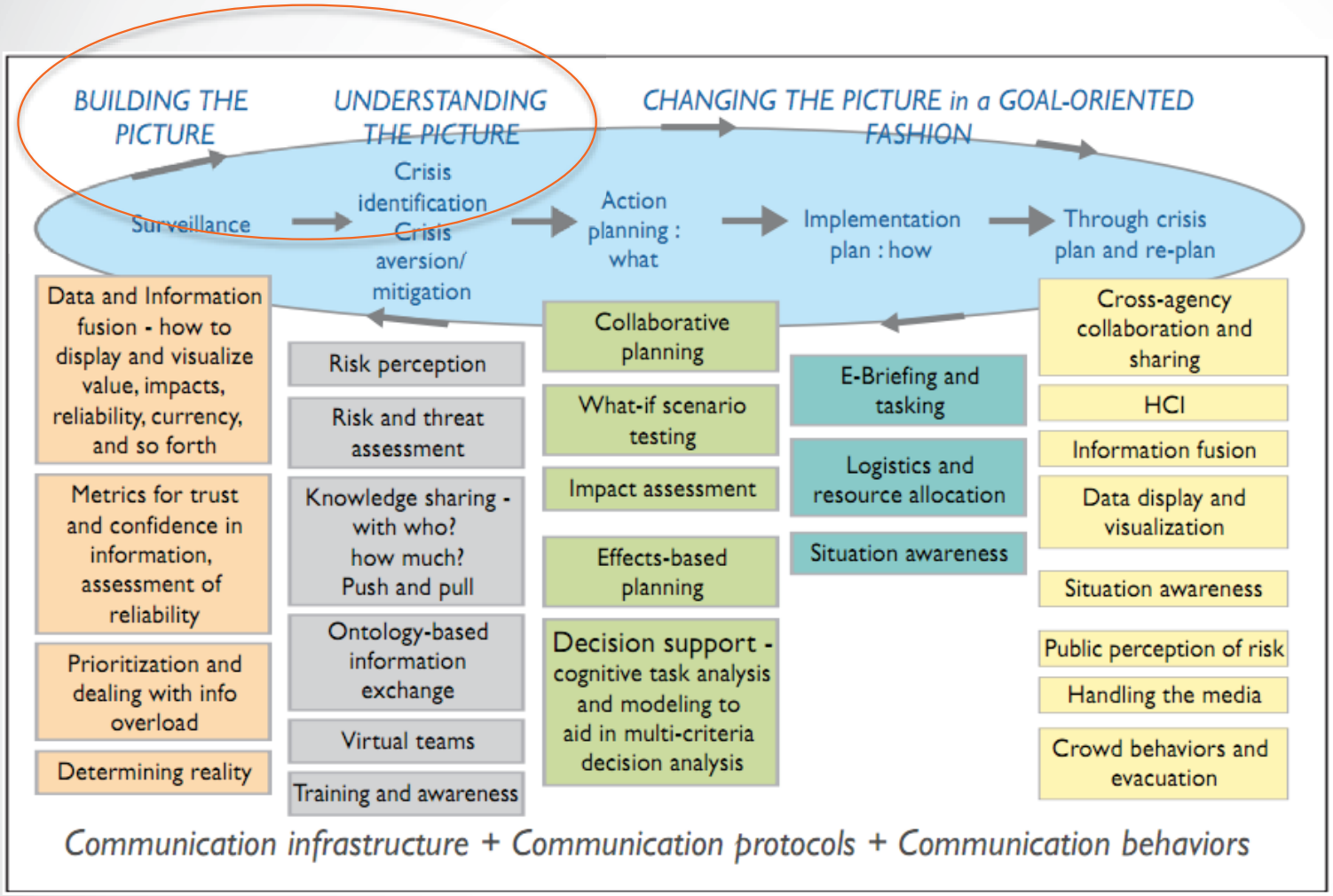
Dusseau, D, J A Negro, and B Clinton. 2001. “Designing User Friendly Situational Awareness Products.” Digital Avionics Systems, 2001. DASC. 20th Conference. doi:10.1109/DASC.2001.963367.

Role of Curated data: Emergency Management Perspective

Leppaniemi, J, P Linna, J Soini, and H Jaakkola. 2009. "Toward a Flexible Service-Oriented Reference Architecture for Situational Awareness Systems in Distributed Disaster Knowledge Management." Management of Engineering & Technology, 2009. PICMET 2009. Portland International Conference on. doi:10.1109/PICMET.2009.5262023.

- Emergency functions are three types:
  - Strategic – that focus on high level coordination of emergency management
  - Tactical – immediate activities in response
  - Operational – management and decision support functions and activities for coordination, prioritization and supporting tactical level functions.
- *Good quality* and a sufficient amount of *timely*, *accurate* and *reliable information* is necessary for successful operations at this level
- Challenges:
  - Comprehensive approach that takes into account Technological, Sociological and Organizational aspects
  - *Types and nature of the information needed*
  - Integration of public and international ad-hoc participants to the knowledge management system

Disaster Response Process



Key needs of those responding to emergencies:

- Absorb information quickly
- Make sense of its meaning, relevance and reliability
- Decide options, actions

Carver, Liz, and Murray Turoff. 2007. "Human-Computer Interaction: The Human and Computer As a Team in Emergency Management Information Systems." Commun. ACM 50 (3) (March): 33–38. doi:10.1145/1226736.1226761. http://doi.acm.org/10.1145/1226736.1226761.

GHRC
Global Hydrology Resource Center

Challenges to: "Building and Understanding the Picture"

- Providing ways to gather accurate and timely information
- Information prioritization- what is useful? what is not?

- Can data and information curation help?
- Can it be automated to filter noise?
- How to best present data/information?

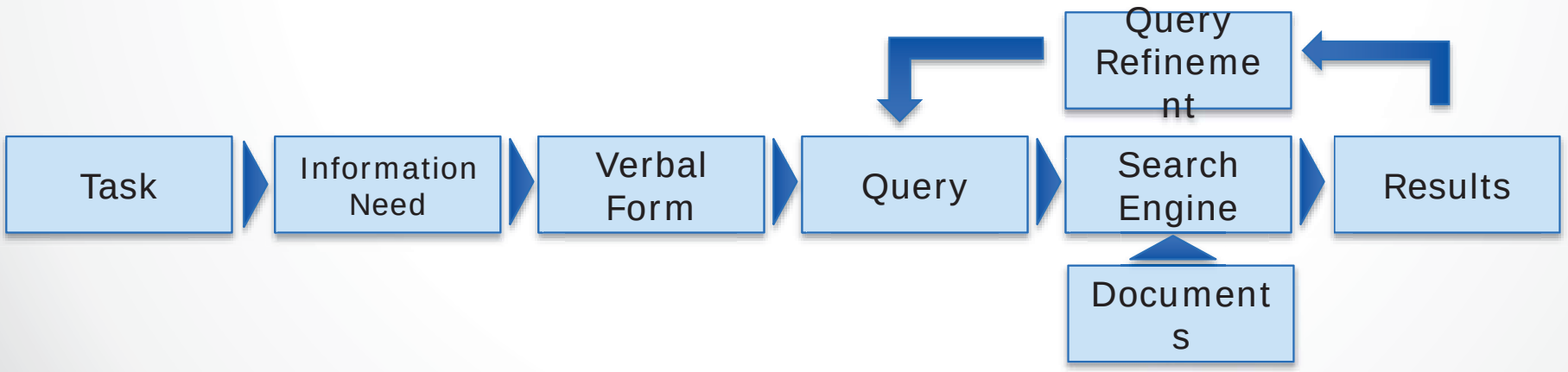
Becomes a specialized search and information display problem

GHRC
Global Hydrology Resource Center

Information Seeking Models

Standard Model

- Identifying the problem (task)
- Articulating the information need
- Formulating the result
- Evaluating the result

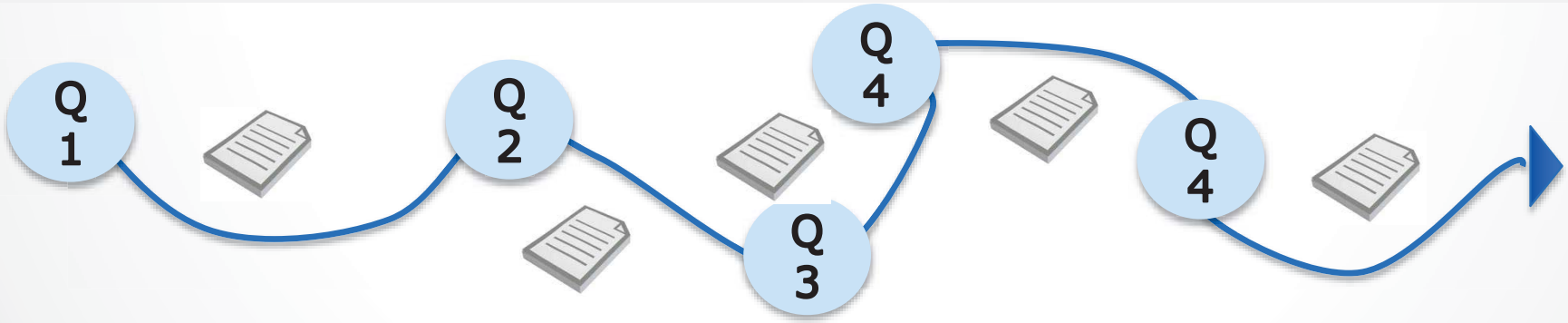


Drawback – assumption that the *task* remains the *same*

Information Seeking Models

Dynamic Model (Bates, 1989)

- User's needs evolve during the process as they interact with the information.
- Analogous to animal foraging for food – learning as they move from one food resource to another
  - Information foraging



Interaction with information leads to new unanticipated goals - "*discovery of latent needs*"

M. J. Bates, "The Design of Browsing and Berry picking Techniques for the Online Search Interface," Online Rev., vol. 13, no. 5, pp. 407–431, 1989

Sensemaking

Behavior generally applied to intelligence analyst and other knowledge workers related to information seeking and use (Bates, 1983)

Framework for sense-making for intelligence analyst (Pirolli and Card, 2005)

- Shoebox – gathering relevant documents into a single collection
- Evidence file – shoebox is curated to be filtered further
- Schema – building a model on how all the information fits
  - Scientific data search – how all the retrieved data fits into the experiment

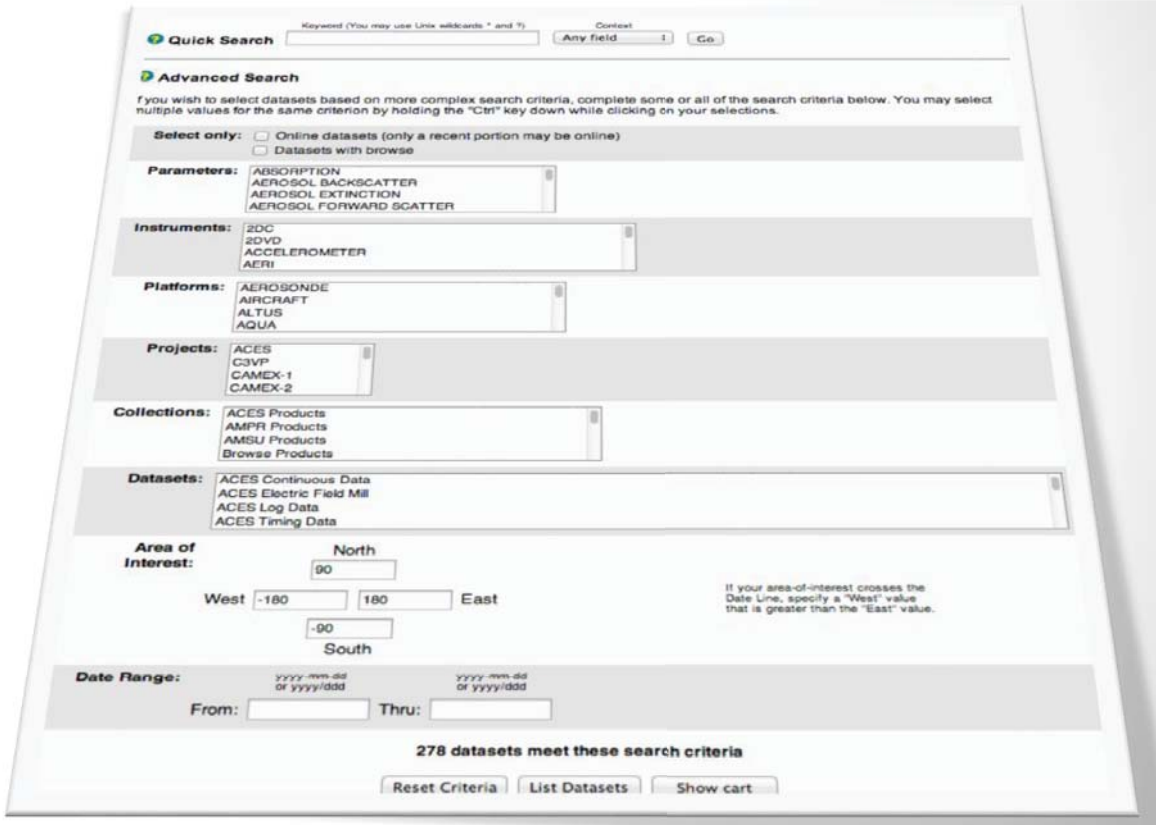
B. Dervin, "An overview of sense-making research: Concepts, methods and results," in Annual Meeting of the International Communication Association, Dallas, TX, 1983

P. Pirolli and S. Card, "The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis 3333 Coyote Hill Road 2 . A Notional Model of Analyst Sense- making," in Proceedings of the 2005 International Conference on Intelligence Analysis, Mclean, VA, 2005.

Data Search Problems: Information Seeking Perspective

Most data search interfaces are built based on the *standard model*

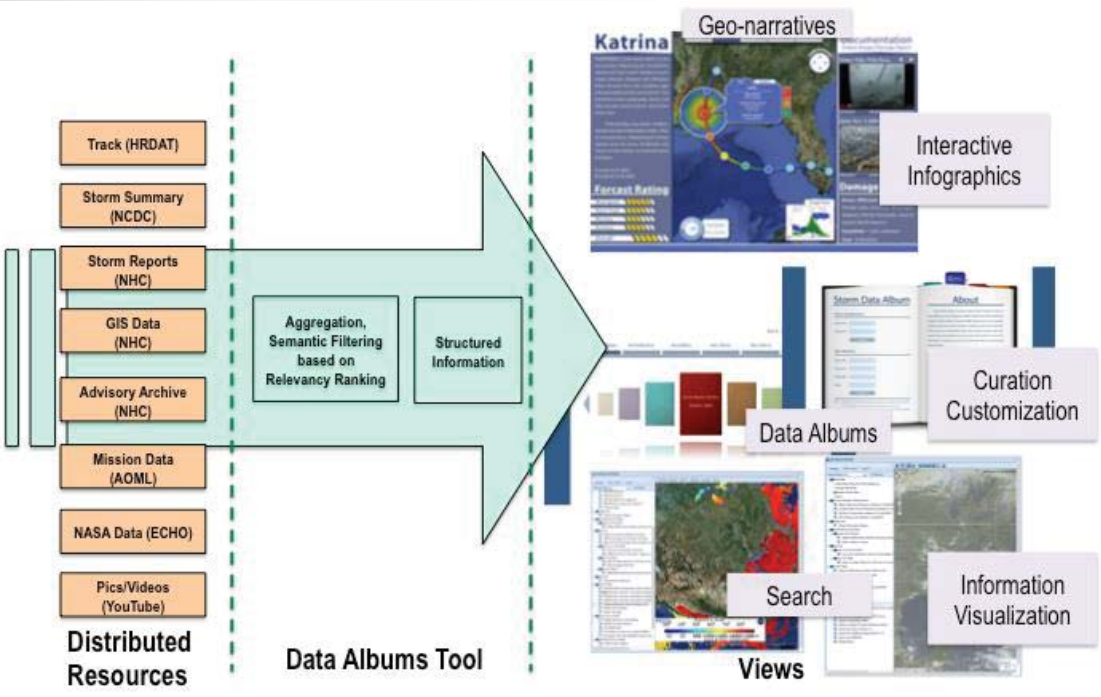
No notion of “discovery of latent needs”



Can we apply the dynamic model and Sense-making framework for data search in Earth Science?

Global Hydrology Resource Center

Data Albums Concept



Aggregated results are presented with visual interactive interfaces to support:

- "discovery of latent needs"
- Dual coding theory – information is best grasped when presented in two modalities

Integrates the "Shoebox" and the "Evidence File" sensemaking concepts

Discovery and aggregation tool that autogenerates "Data Albums"

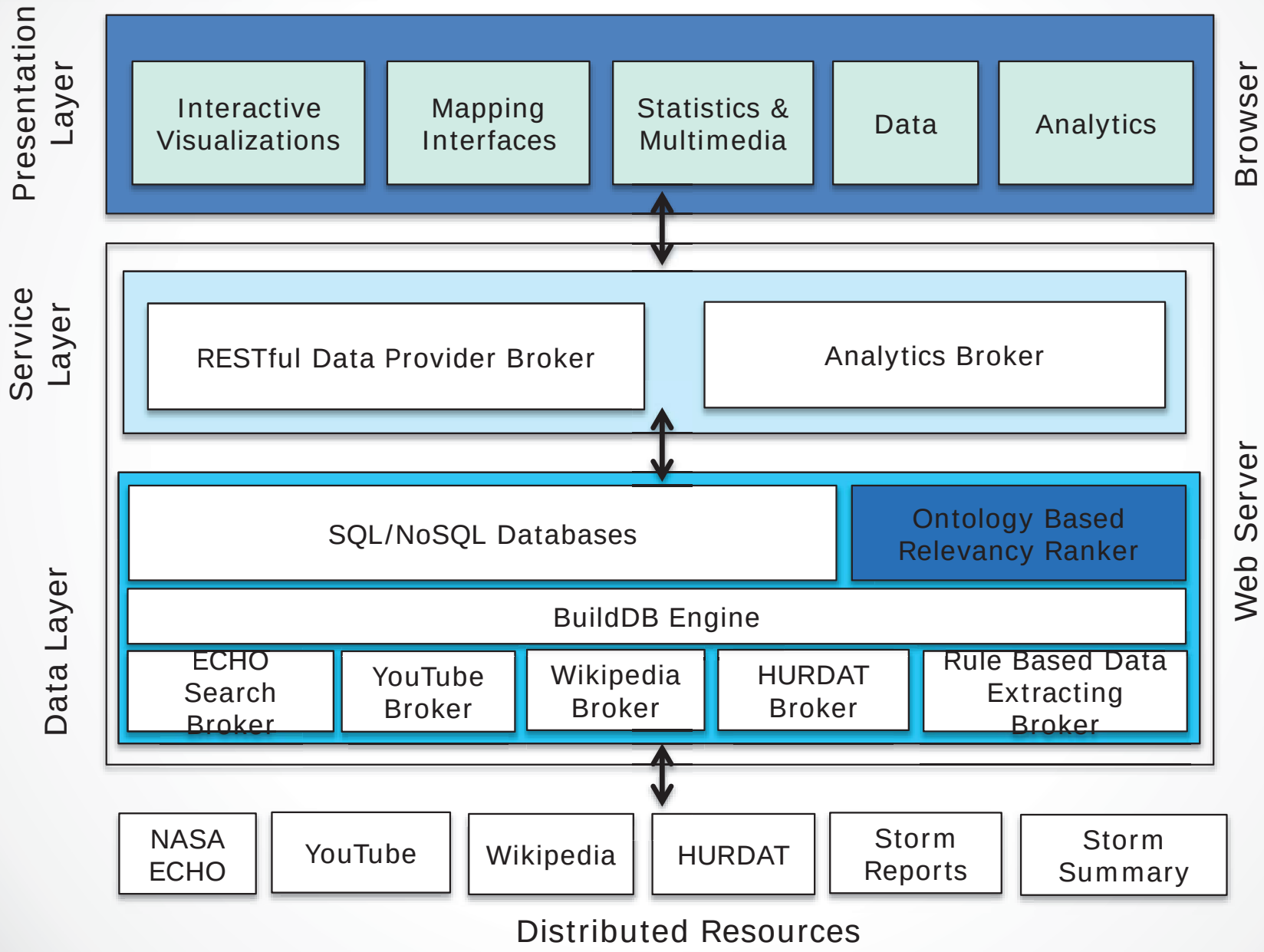
- .. is a compiled collection of data, information, and tools around an event or a theme to support scientific research.

Different curated distributed resources are semantically filtered using an application ontology.

- used both for query expansion and for relevancy ranking text mining algorithm

GHRC
Global Hydrology Resource Center

Architecture



GHRC
Global Hydrology Resource Center

Ontology-based Relevancy Ranking Service

Designed as a general service that can be customized for specific applications.

Utilizes an algorithm that combines ontology based and traditional statistical score to estimate relevancy of a resource (Bouramoul and Kholladi, 2012; Shamsfard et al., 2006)

Relevance Score: Sum of the score of each matched concept obtained by multiplying ontology-based score and the statistical score

$$S_d = \sum_{i=1}^{m} I(c_i) \times tf.idf(c_i, d, D)$$

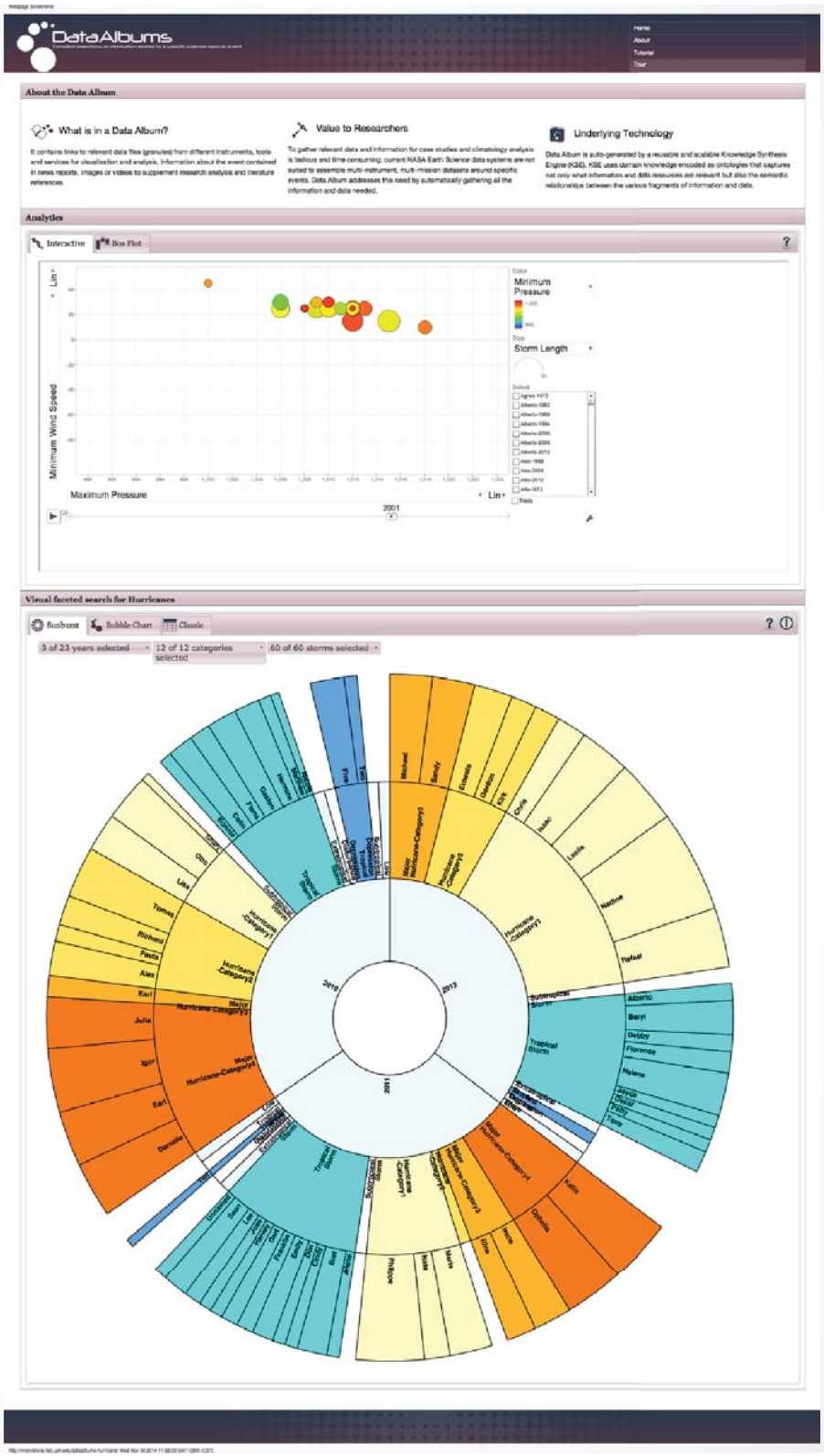
A. Bouramoul and M. Kholladi, "An ontology-based approach for semantic ranking of the web search engines results," in *2012 International Conference on Multimedia Computing and Systems (ICMCS)*, 2012.

M. Shamsfard, A. Nematzadeh, and S. Motiee, "ORank : An Ontology Based System for Ranking Documents," *International Journal of Computer Science*, vol. 1, no. 3, pp. 225–231, 2006.

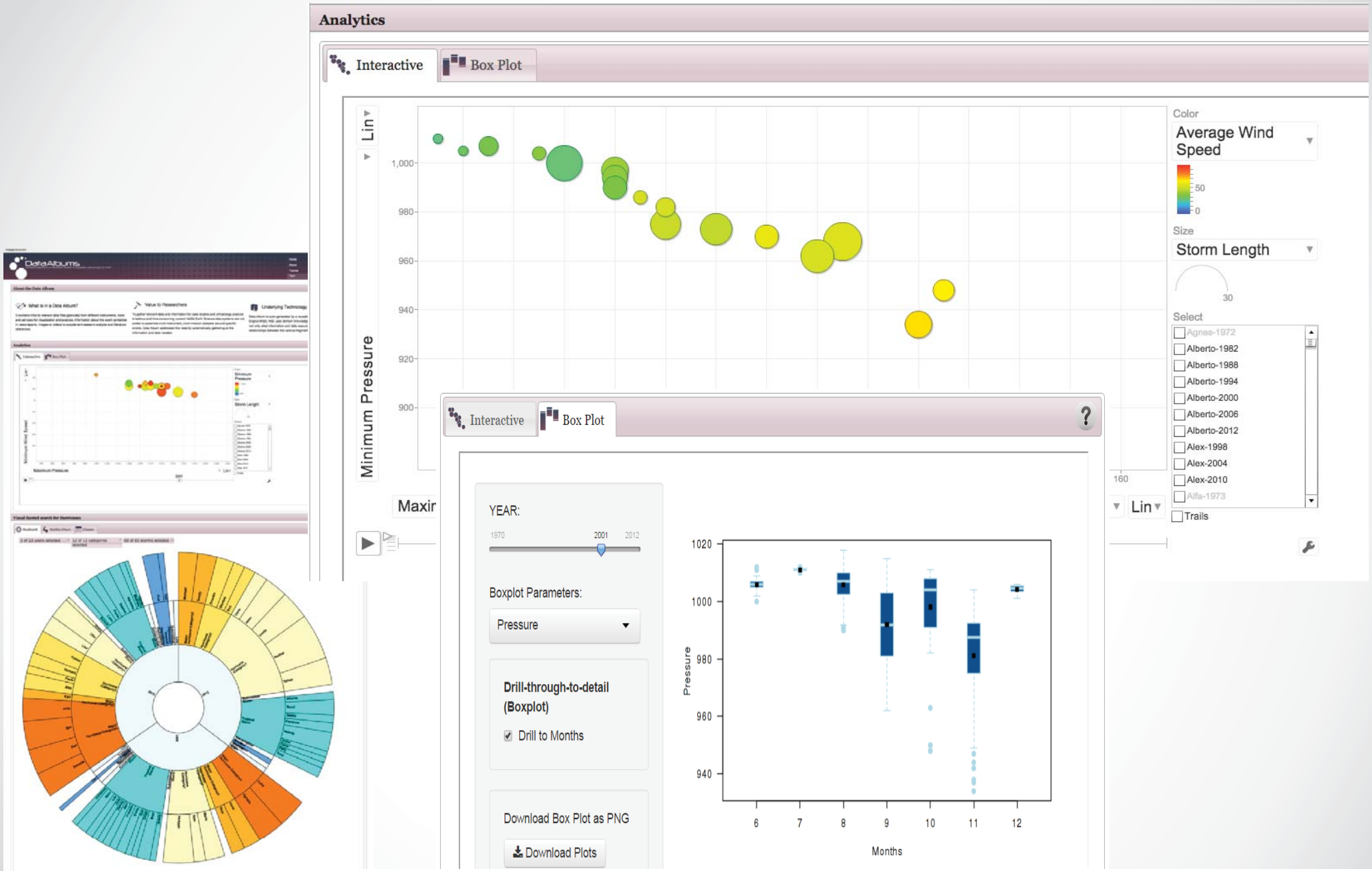
Data Albums Applications: Retrospective Analysis

- Catalog of Hurricane Case Studies at NASA GHRC
- Severe Weather Case study generator at NASA's SPoRT Center

Catalog of Hurricane Case Studies



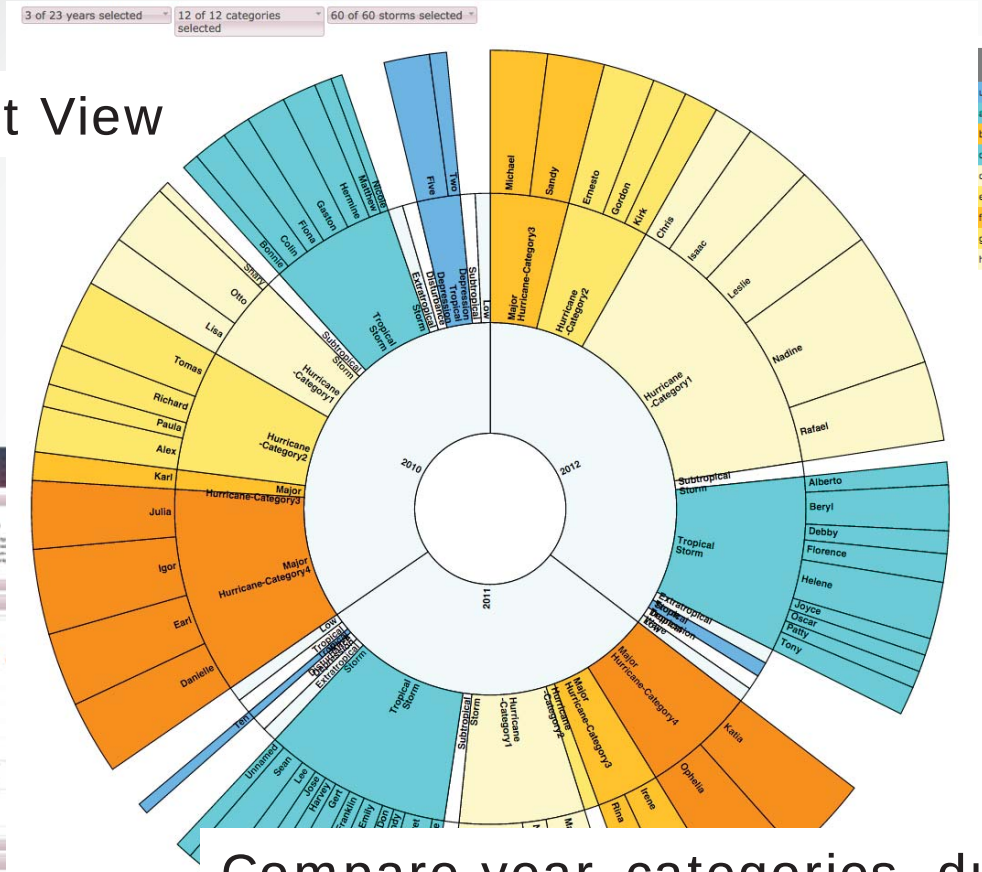
GHRC
Global Hydrology Resource Center

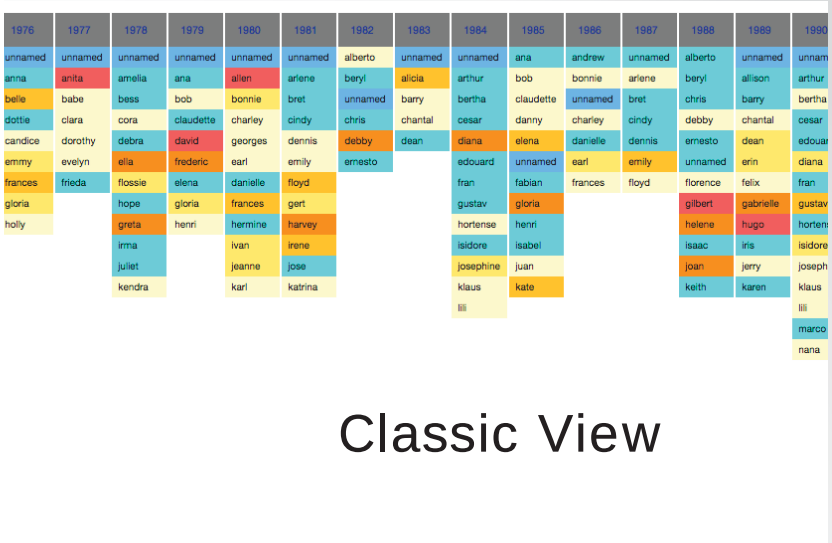


Analytics to explore and find events based on measured parameters

GHRC
Global Hydrology Resource Center

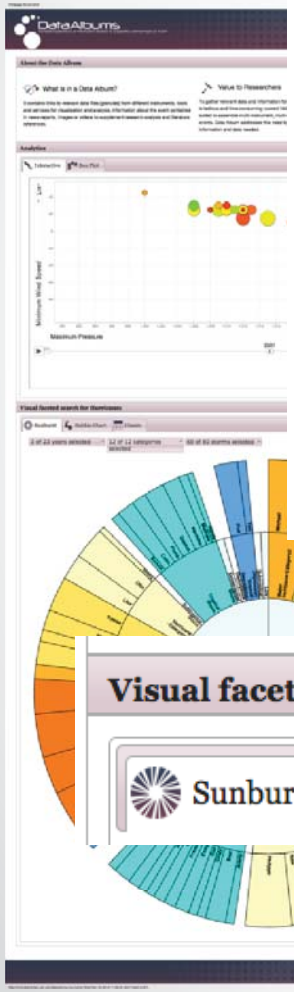
Sunburst View





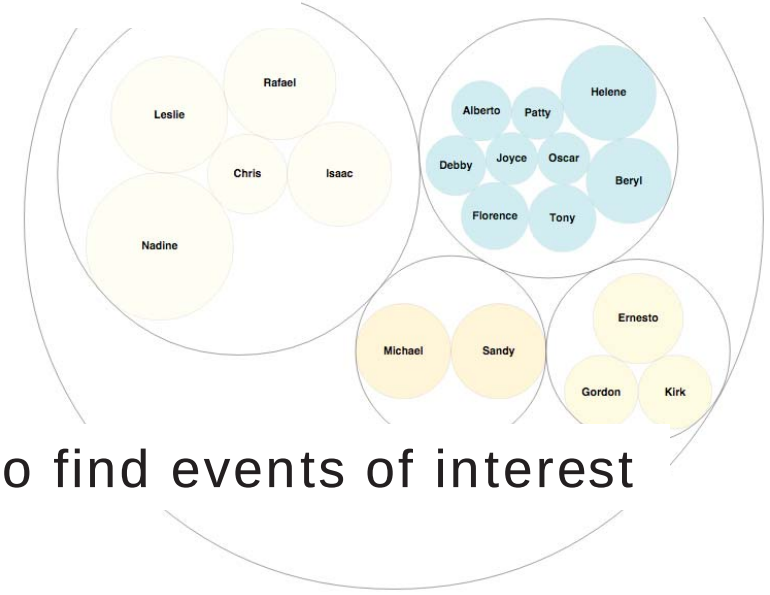
1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
unnamed	unnamed	unnamed	unnamed	unnamed	unnamed	alberto	unnamed	unnamed	ana	andrew	unnamed	alberto	unnamed	unnam
anna	anita	amelia	ana	allen	arlene	beryl	alicia	arthur	bob	bonnie	arlene	beryl	allison	arthur
belle	babe	bess	bob	bonnie	bret	unnamed	barry	bertha	claudette	unnamed	bret	chris	barry	bertha
dottie	clara	cora	claudette	charley	cindy	chris	chantal	cesar	danny	charley	cindy	debby	chantal	cesar
candice	dorothy	debra	david	georges	dennis	debby	dean	diana	elena	danielle	dennis	ernesto	dean	edouar
emmy	evelyn	ella	frederic	earl	emily	ernesto		edouard	unnamed	earl	emily	unnamed	erin	diana
frances	frieda	flossie	elena	danielle	floyd			fran	fabian	frances	floyd	florence	felix	fran
gloria		hope	gloria	frances	gert			gustav	gloria			gilbert	gabrielle	gustav
holly		greta	henri	hermine	harvey			hortense	henri			helene	hugo	horten
		irma		ivan	irene			isidore	isabel			isaac	iris	isidore
		juliet		jeanne	jose			josephine	juan			joan	jerry	joseph
		kendra		karl	katrina			klaus	kate			keith	karen	klaus
								lili						lili
														marco
														nana

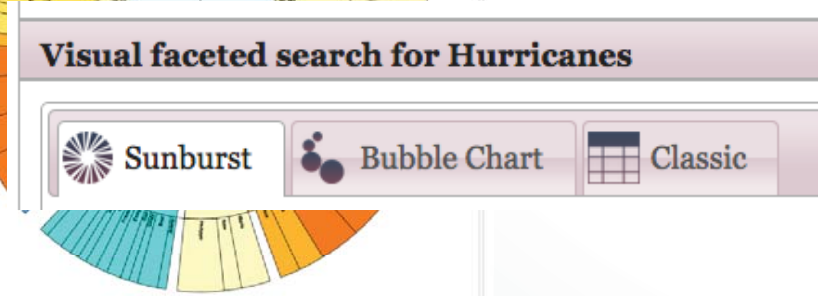
Classic View



Compare year, categories, duration
(Dual encoding theory)

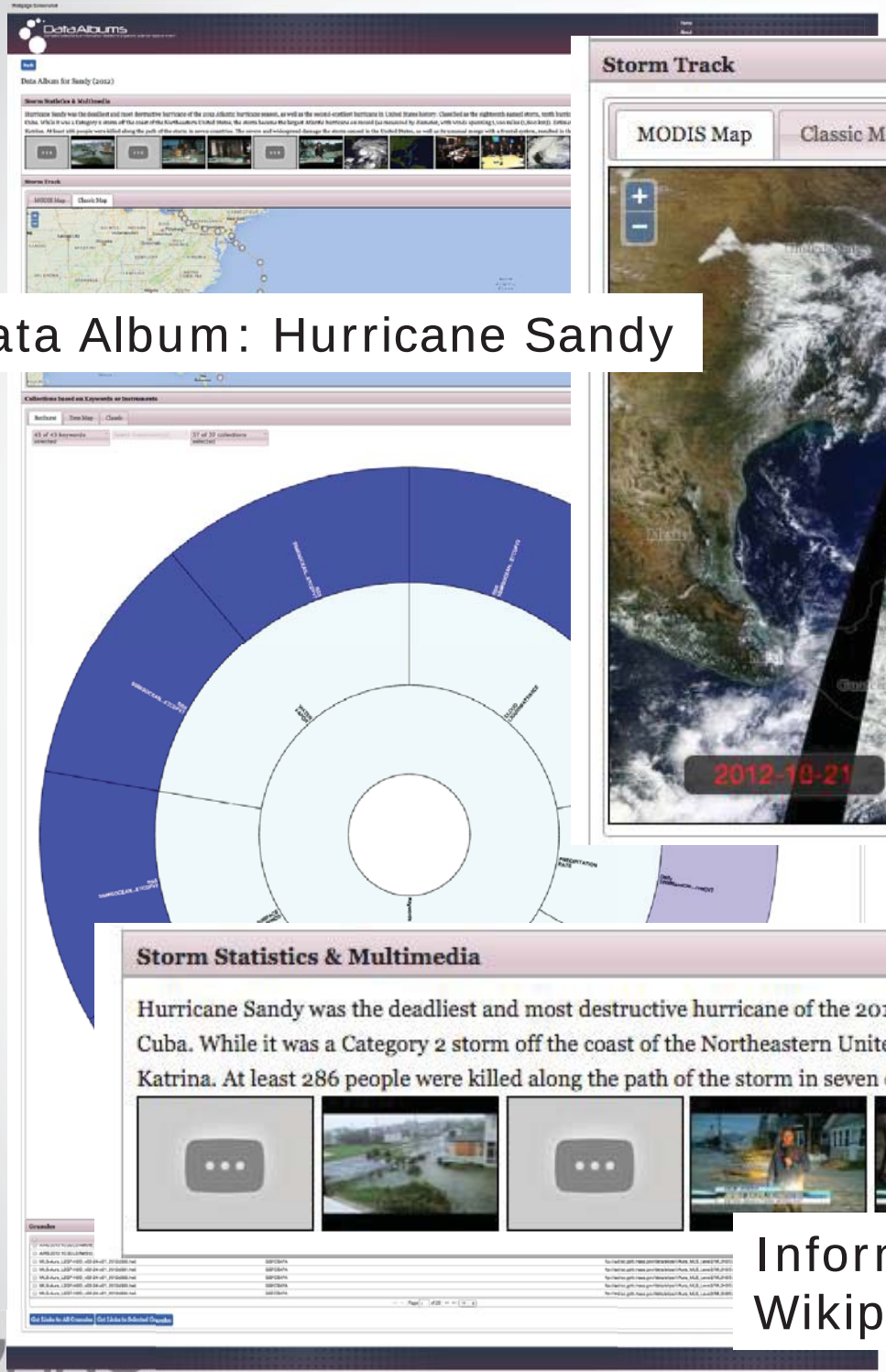
Bubble View





Search the catalog visually to find events of interest

Global Hydrology Resource Center

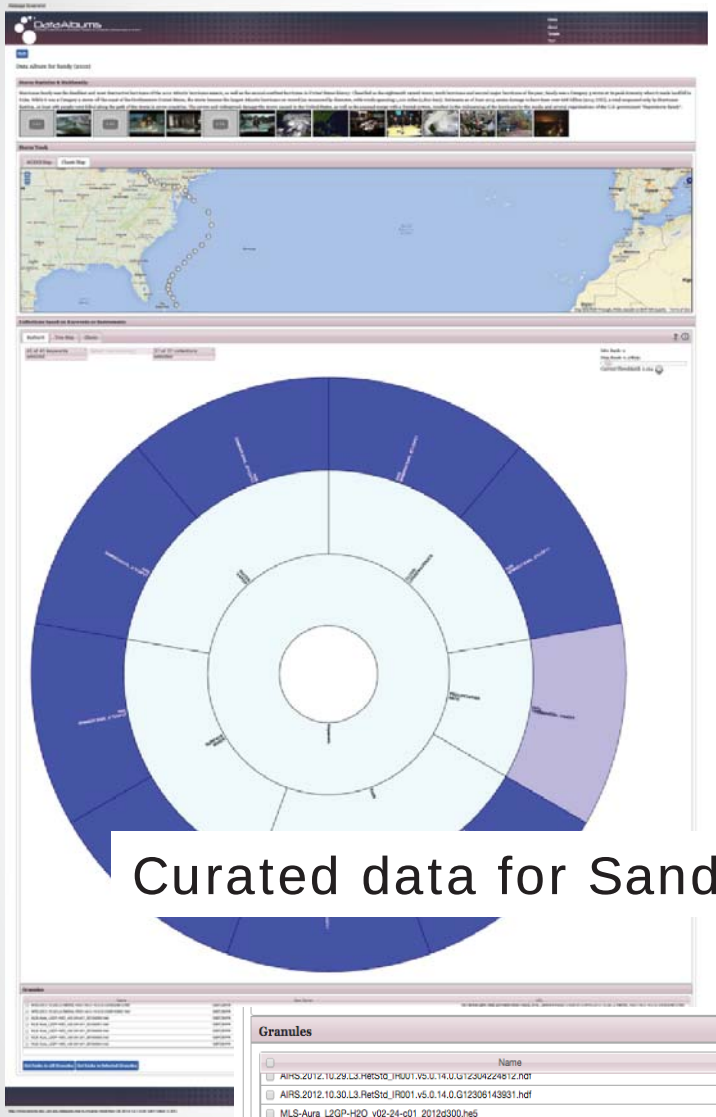


Data Album: Hurricane Sandy

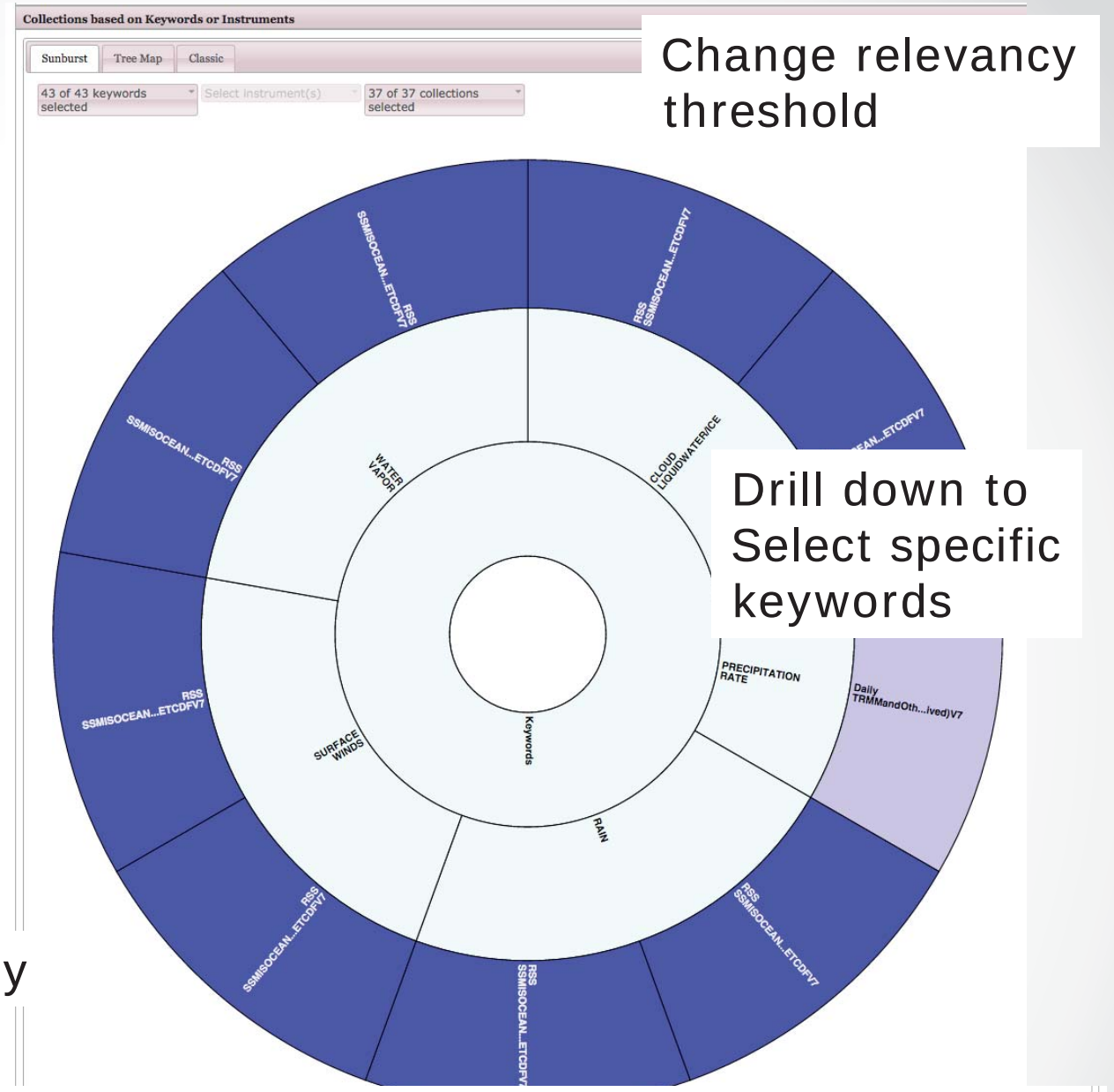
Tracks with MODIS overlay

Information aggregated from the web – Wikipedia, Youtube etc.

Filter the list by keywords, platforms

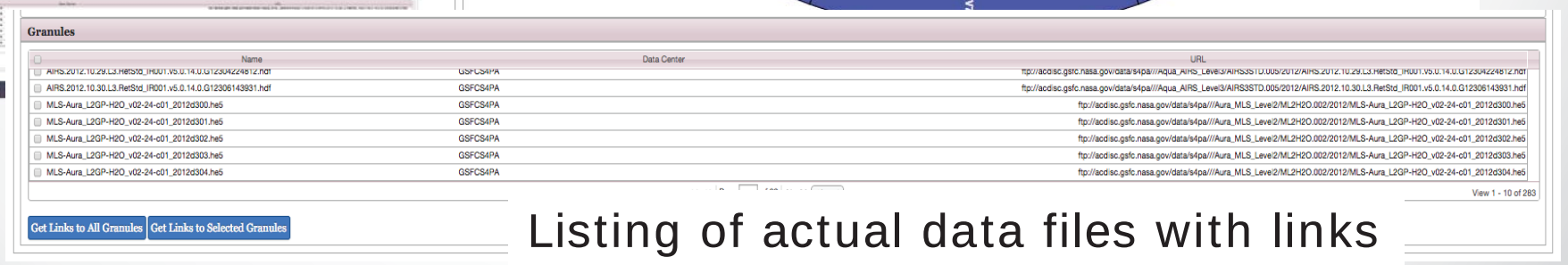


Curated data for Sandy



Change relevancy threshold

Drill down to Select specific keywords



Listing of actual data files with links to download

GHRC
Global Hydrology Resource Center

Severe Weather Case Study Generator for NASA's SPoRT Center

- Center conducts research on unique NASA products and capabilities that can be transitioned to the operational community to solve forecast problems
- Center tests the impact of the NASA datasets on the forecast to ensure and this evaluation process involves use of detailed case studies
- For each case study SPoRT researchers need to search for *forecast texts* from National Weather Service (NWS) describing the atmospheric conditions before and during the storm, *other literature* that may have been written about the event (including news reports, conference papers, or peer-reviewed journal articles), *model input data*, *satellite and in-situ observations, and verification data*—**a time-consuming process that can take days**.

GHRC
Global Hydrology Resource Center

Severe Weather Case Study Generator

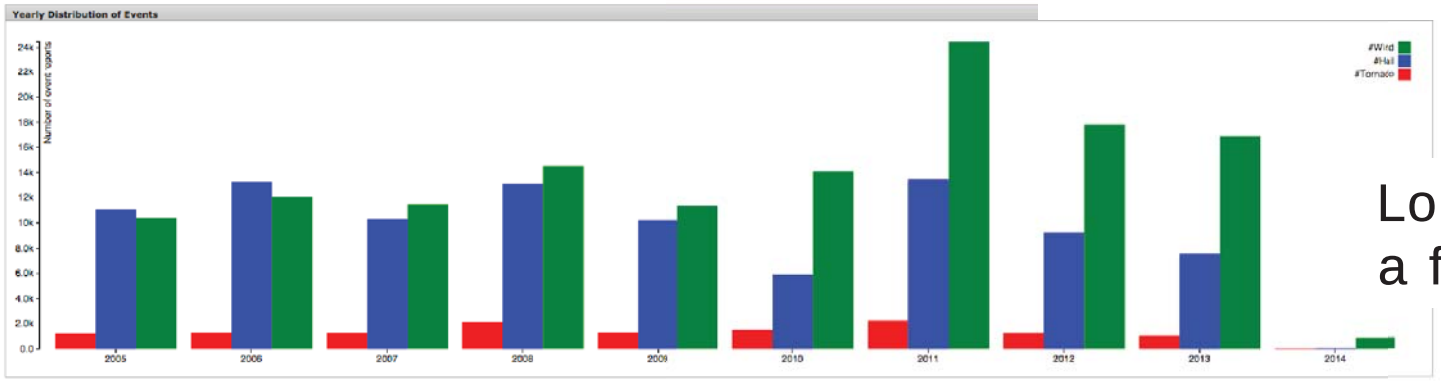
Webpage Screenshot

Data Albums for Severe Storm Case Study

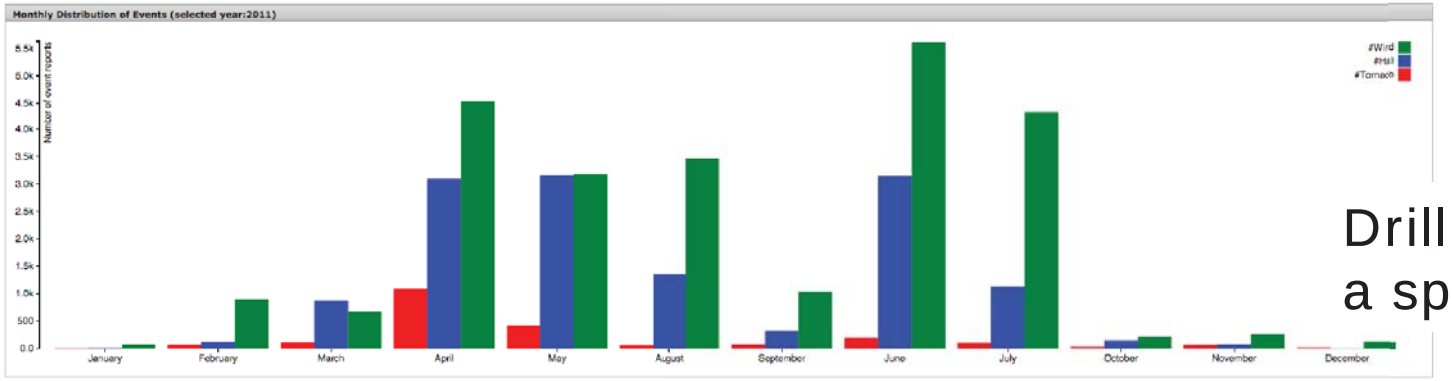
Compiled collections of information for severe storms

Instructions: Following bar charts show the number of events (such as Hail, Wind and Tornado) reported by NOAA Storm Prediction Center (SPC) grouped by year, month and day. In ord
desired year then month and finally a day that you are interested in.

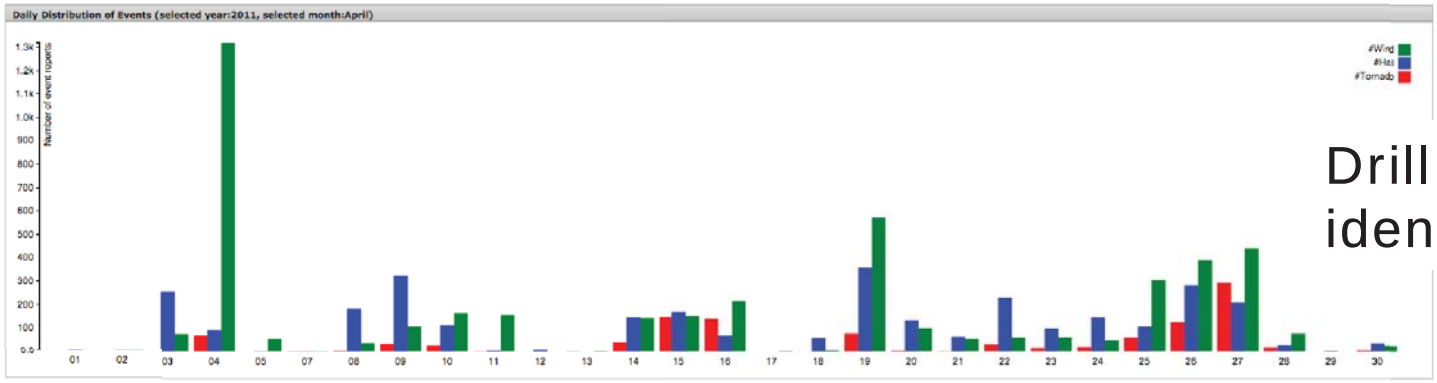
Search for events based on Storm reports – Hail, Wind and Tornado damage



Look at trends over a few years

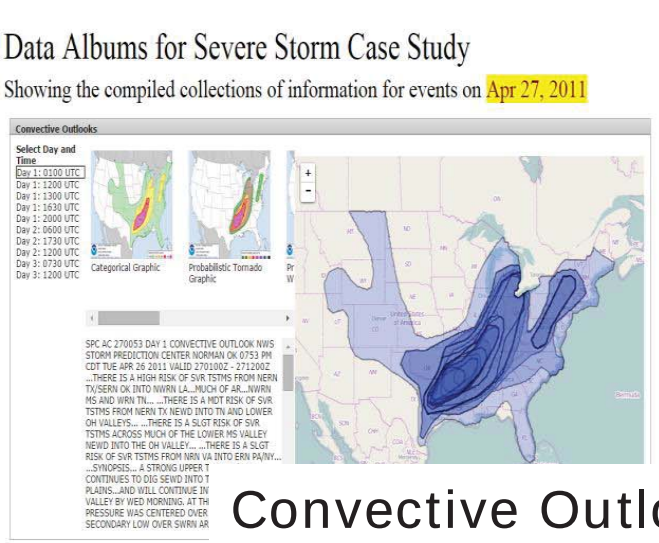


Drill down and look at a specific year



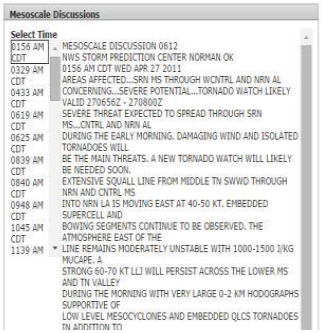
Drill down to a month to identify a day of interest

Data Album: April 27, 2011

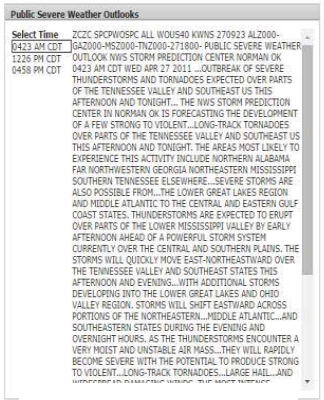


Convective Outlooks

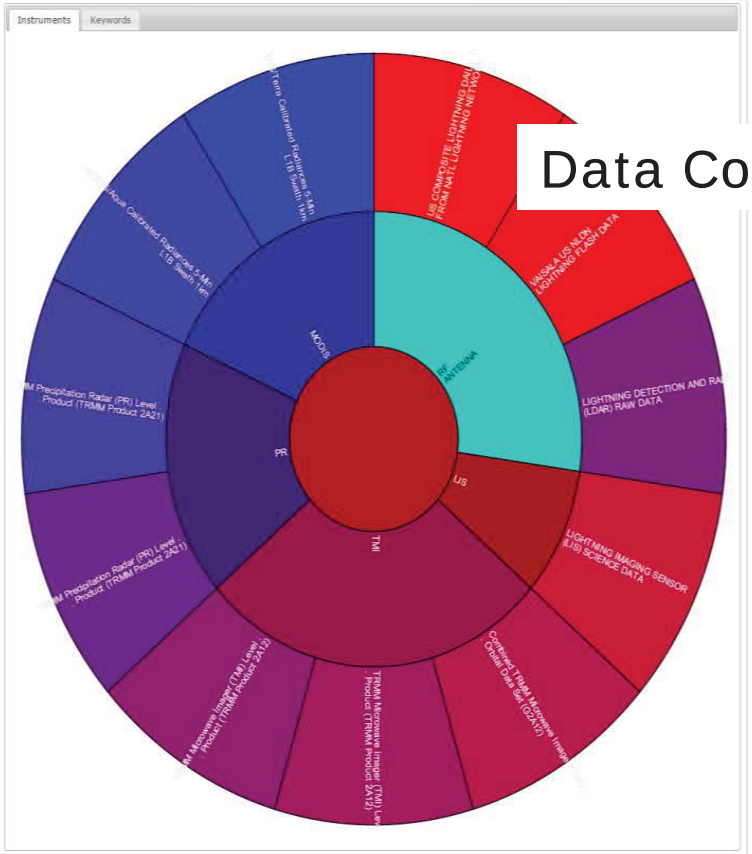
Watches



Mesoscale Discussions

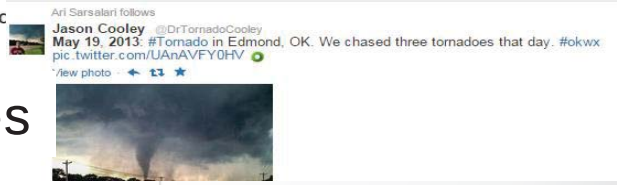


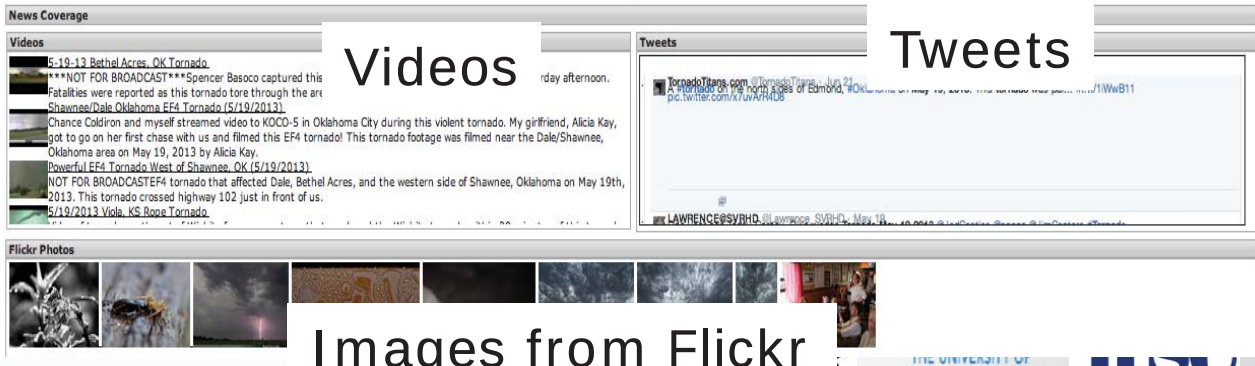
Severe Weather Outlooks



Data Collections

Data Granules





Videos

Tweets

Images from Flickr

Adapting Data Albums Technology

- Retrospective Aggregation to Real time Aggregation
  - Aggregation components to be configured to the right distributed data and information resources
  - Event (disaster) monitoring components
- Improving the existing relevancy ranking algorithm (especially to handle social media feeds)
- Customizations per applications
  - Curated list of resources to monitor and aggregate based on specific event triggers
  - Custom displays for each application or a common API serving a “data album” for an event

Summary

- Data Albums technology can be a useful component within any system designed to support disaster recovery and response
- With the right modifications, this technology can address a critical need – “Building and Understanding the Picture”

Find out more

Dr. Rahul Ramachandran

rahul.ramachandran@nasa.gov

http://innovations.itsc.uah.edu/dataalbums-hurricane/

GHRC
Global Hydrology Resource Center